On Superselection Rules

Domenico Giulini

Max-Planck-Institut für Gravitationsphysik
Potsdam

Universität Mainz, 17. April 2008

Super- and ordinary selection rules

- The notion of **superselection rule** (SSR) was introduced in 1952 by Wick, Wightman, and Wigner in connection with the problem of how to consistently assign intrinsic parity to elementary particles. They understood a SSR as generally expressing "restrictions on the nature and scope of possible measurements".
- The concept of SSR should be contrasted with that of an ordinary **selection rule**. The latter refers to a dynamical inhibition of some transition, usually due to the existence of a conserved quantity, that is, a symmetry. For example, rotational symmetry restricts electric dipole transitions of atoms according to the well known Selection Rule:

$$\Delta J = 0, \pm 1 \quad (\text{except } J = 0 \rightarrow J = 0) \quad \text{and} \quad \Delta M_J = 0, \pm 1$$

Super- and ordinary selection rules (contd.)

- In QM, conserved physical quantities correspond to operators that commute with the Hamiltonian. In contrast, SSRs correspond to quantities that commute with **all** observables.
- Two states, ψ_1 and ψ_2, are separated by a SR if

$$\langle \psi_1 \mid H \mid \psi_2 \rangle = 0$$

  where H is the Hamiltonian. They are separated by a SSR if

$$\langle \psi_1 \mid A \mid \psi_2 \rangle = 0$$

  for **all (physically realisable) observables** A.

Inhibition of superposition principle

- A SSR implies an inhibition to the superposition principle in the sense that the relative phase between ψ_1 and ψ_2 is not measurable and that coherent superpositions of ψ_1 and ψ_2 cannot be verified: Let $\psi_+ = (\psi_1 + \psi_2)/\sqrt{2}$, then

$$\langle \psi_+ \mid A \mid \psi_+ \rangle = \tfrac{1}{2}\big(\langle \psi_1 \mid A \mid \psi_1 \rangle + \langle \psi_2 \mid A \mid \psi_2 \rangle\big) = \mathrm{Tr}(\rho A)$$

  where

$$\rho = \tfrac{1}{2}\big(\mid \psi_1 \rangle\langle \psi_1 \mid + \mid \psi_2 \rangle\langle \psi_2 \mid\big)$$

- Hence, **with respect to the (physically realisable) observables**, the linear combination ψ_+ corresponds to a mixed state. Clearly, this could not be true if all self-adjoint operators in $\mathcal{B}(\mathcal{H})$ corresponded to (physically realisable) observables; e.g. take

$$A = \mid \psi_1 \rangle\langle \psi_2 \mid + \mid \psi_2 \rangle\langle \psi_1 \mid$$

A brief critical intermezzo

- The validity of ordinary selection rules is approximate, in the same way as the corresponding dynamical symmetries are. The latter are always breakable in principle (like e.g. by environmental changes).
- In contrast, a SSR is usually understood as **exact**, i.e. not 'breakable' under all circumstances.
- Is this qualitative distinction well founded? Or is perhaps the 'Super' merely a quantitative statement, like 'very strong'?

Discrete SSR

- There exists a finite or countably infinite family $\{P_i \mid i \in I\}$ of mutually orthogonal and exhaustive projection operators on Hilbert space $\mathcal{H}$:

$$\mathcal{H} = \bigoplus_{i \in I} \mathcal{H}_i \quad \mathcal{H}_i := P_i(\mathcal{H})$$

  such that each observable commutes with all P_i. That is, the **sectors** $\mathcal{H}_i$ reduce the algebra of observables.
- Equivalently, one may also say that states (density matrices) on the given set of observables commute with all P_i, which is equivalent to the identity

$$\rho = \sum_{i \in I} P_i \rho P_i$$

- If $I' = \{i \in I \mid \lambda_i := \mathrm{Tr}(\rho P_i) \neq 0\}$ has more than one element, ρ is mixed:

$$\rho = \sum_{i \in I'} \lambda_i \rho_i \quad \text{where} \quad \rho_i := P_i \rho P_i / \lambda_i$$

Discrete SSR (contd.)

- Only vectors in the **union** $\bigcup_{i \in I} \mathcal{H}_i$ define pure states.
- If, conversely, **any** non-zero vector in this union defines a pure state, with different rays corresponding to different states, one speaks of an **abelian superselection rule**. The $\mathcal{H}_i$ are then called **superselection sectors** or **coherent subspaces** on which the observables act irreducibly.
- The subset $\mathfrak{Z}$ of observables commuting with all observables, called the **centre** of the algebra of physical observables, is then given by

$$\mathfrak{Z} := \left\{ \sum_i a_i P_i \mid a_i \in \mathbb{R} \right\}$$

They are called **superselection-** or **classical observables**.

General SSR

- In general, the Hilbert space will not split into a discrete direct sum of subspaces which are invariant under the action of observables, but rather into a 'continuous direct sum', that is, a direct integral:
$$\mathcal{H} = \int_\Lambda d\mu(\lambda)\mathcal{H}(\lambda)$$
- States are square-integrable functions $f : \lambda \mapsto f(\lambda) \in \mathcal{H}(\lambda)$ and observables are functions $O : \lambda \mapsto O(\lambda) \in \mathcal{B}(\mathcal{H}(\lambda))$. Closed subspaces of $\mathcal{H}$ which are left invariant by the observables are precisely given by
$$\mathcal{H}(\Delta) = \int_\Delta d\mu(\lambda)\mathcal{H}(\lambda)$$
where $\Delta \subset \Lambda$ is any measurable subset of non-zero measure. In general, a single $\mathcal{H}(\lambda)$ will not be a subspace (unless the measure has discrete support at λ).
- States are separated by a SSR if their supports in Λ are disjoint. Hence one also speaks of **disjoint states**.

SSRs discussed in the literature

In the literature, SSRs are discussed in connection with a variety of superselection-observables, most notably

- univalence
- overall mass (in non-relativistic QM),
- electric charge
- baryonic and leptonic charge
- time

Algebraic Theory

- In Algebraic Quantum Mechanics, a system in isolation is characterised by an abstract C^*–algebra $\mathcal{C}$.
- Depending on contextual physical conditions, one chooses a faithful representation $\pi : \mathcal{C} \rightarrow \mathcal{B}(\mathcal{H})$ in the (von Neumann) algebra of bounded operators on some Hilbert space $\mathcal{H}$. After completing the image of π in the weak operator-topology on $\mathcal{B}(\mathcal{H})$ (dressing) one obtains a von Neumann sub-algebra $\mathfrak{M} \subset \mathcal{B}(\mathcal{H})$, called the **algebra of (bounded) observables**. The physical observables proper correspond to the self-adjoint elements of $\mathfrak{M}$.
- SSRs are now said to exists if the commutant

$$\mathfrak{M}' := \{B \in \mathcal{B}(\mathcal{H}) \mid [A, B] = 0 \,\forall A \in \mathfrak{M}\}$$

is not trivial (different from multiples of the unit). Note that this makes reference to the Hilbert space $\mathcal{H}$ and is hence not intrinsic to $\mathcal{C}$. Sectors are defined by the projectors in $\mathfrak{M}'$. Abelian SSRs are characterised by $\mathfrak{M}'$ being abelian. Note also that $\mathfrak{M}'' = \mathfrak{M}$ holds for any von Neumann algebra $\mathfrak{M} \subseteq \mathcal{B}(\mathcal{H})$.

Algebraic Theory (contd.)

- Dirac's requirement of the existence of a **maximal commuting set of observables** now reads as follows: There exists a maximal abelian subalgebra $\mathfrak{A} \subset \mathfrak{M}$. This is a non-trivial requirement since maximality refers to $\mathcal{B}(\mathcal{H})$ (not to $\mathfrak{M}$).

$$\mathfrak{A} \text{ max. abelian} \Leftrightarrow \mathfrak{A} = \mathfrak{A}' \begin{cases} \mathfrak{A} \subseteq \mathfrak{A}' & \mathfrak{A} \text{ is abelian} \\ \mathfrak{A} \supseteq \mathfrak{A}' & \mathfrak{A} \text{ is maximal} \end{cases}$$

- **Theorem:** This is equivalent to the existence of a cyclic vector.
- **Theorem:** This is equivalent to $\mathfrak{M}'$ being abelian.
- **Proof of the latter:** If $\mathfrak{A} \subseteq \mathfrak{M}$ is max. abelian, then $\mathfrak{M}' \subseteq \mathfrak{A}' = \mathfrak{A} \subseteq \mathfrak{M}$ and hence $\mathfrak{M}' \subseteq \mathfrak{M}''$, i.e. $\mathfrak{M}'$ is abelian. Conversely, let

$$\mathfrak{M}' \subseteq \mathfrak{M} \qquad (\mathfrak{M} \text{ abelian}) \tag{1}$$

$$\mathfrak{A} = \mathfrak{A}' \cap \mathfrak{M} \qquad (\mathfrak{A} \text{ max. abelian in } \mathfrak{M}) \tag{2}$$

Since $\mathfrak{A} \subseteq \mathfrak{M}$ implies $\mathfrak{M}' \subseteq \mathfrak{A}'$, we have

$$\mathfrak{M}' \stackrel{(1)}{=} \mathfrak{M} \cap \mathfrak{M}' \subseteq \mathfrak{M} \cap \mathfrak{A}' \stackrel{(2)}{=} \mathfrak{A}$$

Now, $\mathfrak{M}' \subseteq \mathfrak{A}$ implies $\mathfrak{A}' \subseteq \mathfrak{M}$ so that (2) implies $\mathfrak{A} = \mathfrak{A}'$.

SSR and Decoherence

- Decoherence explains the invisibility of certain states due to the ubiquitous dynamical process of **phase dislocalisation**.
- Formally, the process is of the following form, which makes the connection with SSR obvious:

$$\rho \equiv \sum_{i,j} P_i \, \rho \, P_j \quad \rightsquigarrow \quad \sum_i P_i \, \rho \, P_i$$

- The complete set of projection operators, $\{P_i\}$, depends on the interaction of the system with the environment. They define the sectors across which phase relations cannot be measured locally ($\rightarrow$ 'pointer basis').
- This gives rise to the notion of **environmentally induced SSRs**.
- Traditional arguments for SSRs suggest a more fundamental nature. These will be discussed next.

SSR and conserved additive quantities

- Let Q be the operator corresponding to a ‘charge-like’ quantity. This means that it is 1) conserved under time evolution , 2) additive under composition of systems, and 3) for subsystems independent of the state of the complementary system. If $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}$ then $Q = Q_1 \otimes 1 + 1 \otimes Q_2$.
- It is not difficult to show that superpositions of Q eigenstates cannot form through the processes of time evolution, composition, and decomposition of systems.
- Exact von Neumann measurements respecting the conservation of Q are impossible for operators not commuting with Q (Wigner 1952). Approximate measurements are only possible to the extend that the apparatus can be prepared in a superposition of Q eigenstates (Araki & Yanase 1960), in which case the total system (laboratory) cannot be in an eigenstate of Q.
- This seems to give rise to an abundance of SSRs.

Intermezzo: Proof of Wigner's theorem

- Let S be the system to be measured, A the measuring apparatus. Then $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_A$. The charge-like quantity is represented by the operator $Q = Q_S \otimes 1 + 1 \otimes Q_A$, the observable of S by $P \in \mathcal{B}(\mathcal{H}_S)$.
- Let $\{|s_n\rangle\} \subset \mathcal{H}_S$ be a set of normalised eigenstates for P so that $P|s_n\rangle = p_n|s_n\rangle$. Let $U \in \mathcal{B}(\mathcal{H})$ be the unitary evolution operator for the von Neumann measurement and $\{|a_n\rangle\} \subset \mathcal{H}_A$ a set of normalised 'pointer states' with neutral pointer-position a_0, so that

$$U\big(|s_n\rangle|a_0\rangle\big) = |s_n\rangle|a_n\rangle$$

- We assume the total Q to be conserved during the measurement, i.e. $[U, Q] = 0$. Clearly $\langle a_n \mid a_m\rangle \neq 1$, for, otherwise, this is not a measurement, since $\langle a_n \mid a_m\rangle = 1$ if and only if $|a_n\rangle = |a_n\rangle$). Then

$$\begin{aligned}(p_n - p_m)\langle s_n|Q_S|s_m\rangle &= \langle s_n|\ [P, Q_S]\ |s_m\rangle = \langle s_n|\langle a_0|\ [P \otimes 1, Q]\ |s_m\rangle|a_0\rangle \\ &= (p_n - p_m)\langle s_n|\langle a_0|\ U^\dagger QU\ |s_m\rangle|a_0\rangle \\ &= (p_n - p_m)\langle s_n|\langle a_n|\ Q_S \otimes 1 + 1 \otimes Q_A\ |s_m\rangle|a_m\rangle \\ &= \langle a_n|a_m\rangle\ (p_n - p_m)\langle s_n|Q_S|s_m\rangle\end{aligned}$$

$$\Leftrightarrow \quad \langle s_n|Q_S|s_m\rangle = 0 \text{ if } p_n \neq p_m \quad \Leftrightarrow \quad [Q_S, P] = 0$$

SSR and conserved additive quantities (contd.)

- Clearly, there are many conserved quantities without associated SSR, like momentum, angular momentum, etc. One crucial observation here is that these quantities are physically always understood as **relative** to a system of reference that, by its very definition, must have certain localisation properties which exclude the total system to be in eigenstates of **relative** (linear and angular) momenta.
- "...SSRs do not exist. The belief of 3W is wrong because their basic assumption about physics is incorrect. [...] Different observers can disagree on whether two states are coherently superposed" (Mirman 1977; Aharonov & Susskind 1967)
- The question may be asked whether there is a fundamental difference in this respect between the Noether charges for spacetime symmetries (momentum, angular momentum etc.) and those for global gauge transformations, like electric charge. (Aharonov & Susskind 1967 versus 3W 1970)

SSR and symmetries

- Certain symmetry considerations seem to unambiguously prove the existence of SSR. These proofs rely on the fact that in QM and QFT it is sufficient to implement symmetries by (anti-)unitary **ray** representations:

$$U(g_1)U(g_2) = \omega(g_1, g_2)\, U(g_1 g_2)$$

where $\omega : G \times G \to \mathrm{U}(1) := \{z \in \mathbb{C} \mid |z| = 1\}$ is the so-called **multiplier** that satisfies (associativity of group action)

$$\omega(g_1, g_2)\omega(g_1 g_2, g_3) = \omega(g_1, g_2 g_3)\omega(g_2, g_3)$$

- Any function $\alpha : G \to \mathrm{U}(1)$ allows to redefine $U \mapsto U'$ via $U'(g) := \alpha(g)U(g)$, which amounts to a redefinition $\omega \mapsto \omega'$:

$$\omega'(g_1, g_2) = \frac{\alpha(g_1)\alpha(g_2)}{\alpha(g_1 g_2)}\, \omega(g_1, g_2)$$

- Two multipliers ω and ω' are called **similar** if this holds for some function α. A multiplier is called **trivial** if it is similar to $\omega \equiv 1$, in which case the ray-representation is a proper representation in disguise.

SSR and symmetries (contd.)

- **Theorem:** Given unitary ray-representations $U_{1,2}$ of G on $\mathcal{H}_{1,2}$, respectively, with non-similar multipliers $\omega_{1,2}$, then no ray-representation of G on $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ exists which restricts to $U_{1,2}$ on $\mathcal{H}_{1,2}$ respectively.
- From this a SSR follows from the **hypothesis of a definite symmetry group** and Wigner's theorem, according to which symmetries are always implementable by (anti-) unitary ray-representations. The largest subset of rays in $\mathcal{H}_1 \oplus \mathcal{H}_1$ on which G acts is the set of rays in $\mathcal{H}_1 \cup \mathcal{H}_1$.
- An example is given by the SSR of univalence, that is, between states of integer and half-integer spin. Here G is the group $SO(3)$ of proper spatial rotations. For integer spin it is represented by proper unitary representations, for half integer spin with non-trivial multipliers.

Example: Univalence

- We need to prove that multipliers of spin-1/2 representation are not trivial. This is implied if they are not trivial after restriction to a subgroup $G \subset SO(3)$. Choose for G the four-element abelian group D_4 (isomorphic to $\cong \mathbb{Z}_2 \times \mathbb{Z}_2$), given by the identity and the 180° rotations g_i ($i = 1, 2, 3$) about the x, y, z axes.
- Have $\omega(g_a, g_a) = -1$ and $\omega(g_a, g_b) = -\omega(g_b, g_a) = 1$ for b cyclic successor of a.
- Since G is abelian, a redefinition

$$\omega'(g_1, g_2) = \underbrace{\frac{\alpha(g_1)\alpha(g_2)}{\alpha(g_1 g_2)}}_{D(g_1, g_2)} \omega(g_1, g_2)$$

  amounts to a symmetric correction, $D(g_1, g_2) = D(g_2, g_1)$, which cannot cancel the antisymmetric ω.
- The virtue of this proof is that it makes clear that no smoothness assumptions concerning the function α need to be made.

Caveats

- The symmetry argument crucially depends on the 'size' of the proposed symmetry group G. In case of univalence, no SSR would result if $SU(2)$ instead of $SO(3)$ were required to act on state space (cf. S. Weinberg, QTF I).
- Generally, a ray representation of G with non-trivial multiplier ω can always be considered as proper representation of a central extension G_ω of G by $U(1)$, given by (multiplicative structure on the set $S^1 \times G$)

$$(z_1, g_1)(z_2, g_2) = (z_1 z_2 \omega(g_1, g_2)\,,\; g_1 g_2)$$

  The proper representation U_ω of G_ω follows form the ray representation U of G via

$$U_\omega\big((z, g)\big) := z \cdot U(g)$$

- Independently of ω, there exist **universal central extensions** $\hat{G}$ of G by abelian groups A, so that **any** ray representation of G corresponds to a proper representation of $\hat{G}$.

Example: Mass superselection

- Consider ordinary QM for N particles with masses m_a, interacting via a Galilei invariant potential $V(\|\vec{x}_a - \vec{x}_b\|)$. The inhomogeneous Galilei group acts as symmetries of the Schrödinger equation via a unitary ray representation with non-trivial multiplier (where $M = \sum_a m_a$ and $g = (R, \vec{v}, \vec{a}, b)$):

$$\omega(g_1, g_2) = \exp\left\{ i\,\frac{M}{\hbar}\left(\vec{v}_1 \cdot R_1 \cdot \vec{a}_2 + \tfrac{1}{2}\vec{v}_1^2 b_2\right)\right\}$$

- For different overall masses M and M' the multipliers are inequivalent. Hence a **SSR for mass** results ('Bargmann SSR').
- But: mass superselection for what system? In order to makes sense of that, **mass must be a dynamical variable**.

Example: Mass superselection (contd.)

- Adjoin N new pairs of conjugate variables (λ_a, m_a) and consider minimally extended action

$$S[\lambda_a, \vec{x}_a\,; m_a, \vec{p}_a] = \int dt \left\{ \sum_a m_a \dot{\lambda}_a + \vec{p}_a \cdot \dot{\vec{x}}_a - H(m_a, \vec{x}_a, \vec{p}_a) \right\}$$

- Hamilton's equations give
  - $\dot{m}_a = 0 \Rightarrow m_a = \text{const.}$
  - $\dot{\vec{x}}_a = \partial H / \partial \vec{p}_a \,, \quad \dot{\vec{p}}_a = -\partial H / \partial \vec{x}_a$ (as before)
  - $\dot{\lambda}_a = \partial V / \partial m_a - \vec{p}_a^{\,2} / 2 m_a^2$ (by quadrature)
- The symmetry group of this system is the 11-dimensional Schrödinger group (central $\mathbb{R}$-extension of Galilei), which does not give rise to any SSR (D.G. 1995).
- Mass superselection corresponds to removing the average λ position, $\sum_a \lambda_a$, from the observables, thereby creating a non-trivial centre of the algebra of observables.

SSR and locality & causality

- In QFT SSRs emerge through the requirement of locality and causality: $\mathfrak{M}$ is given by the (weak closure of) observables localised in space and time (smearing functions of compact support). Two localised observables commute if their space-time support is causally disconnected (spacelike separated).
- SSRs then arise for charge-like quantities which can be measured on spheres of arbitrarily large spatial radii due to Gauß' law $\rho = \vec{\nabla} \cdot \vec{E}$:

$$\langle \Psi \mid [A, Q] \mid \Phi \rangle = \lim_{R \to \infty} \Big\langle \Psi \Big| \int_{\|\vec{x}\| \leq R} d^3x \quad \big[A, \rho(\vec{x}) \big] \quad \Big| \Phi \Big\rangle$$

$$= \lim_{R \to \infty} \Big\langle \Psi \Big| \int_{\|\vec{x}\| = R} d^2\sigma \quad \big[A, \vec{E}(\vec{x}) \big] \cdot \vec{n} \quad \Big| \Phi \Big\rangle$$

- This vanishes if the 2-sphere $\|\vec{x}\| = R$ is spacelike separated from A's support. Gauß's law can be justified as operator identity (Strocchi & Wightman 1974).

SSR and locality & causality (contd.)

- The foregoing argument seems to suggest an abundance of SSRs in field theory, one for each Gauß-like law.
- For example, in General Relativity, the Poincaré charges **mass, linear-, and angular momentum** are all given by surface integrals over 2-spheres at spacelike infinity:

$$m = \lim_{R\to\infty} \left\{ \frac{c^2}{16\pi G} \int_{S_R} d^2\sigma \; n^a \left(\partial_b g_{ab} - \partial_a g_{bb} \right) \right\}$$

$$p_\xi = \lim_{R\to\infty} \left\{ \frac{c^2}{8\pi G} \int_{S_R} d^2\sigma \; n^a \left(K_{ab} - \delta_{ab} K_{cc} \right) \xi^b \right\}$$

- In this context the restriction to local observables seems less well justified. For example, an observable not commuting with angular momentum would be the spatial orientation relative to a background reference frame by “looking at fixed stars” ($\rightarrow$ *extra-galactic celestial reference frame*).

Asymptotic DOF

- Consider electromagnetic system spatially enclosed within sphere $r \leq R$, so that normal component of $\vec{j}$ and tangential components of $\vec{B}$ vanish on the boundary.
- Consistent variational principle requires addition of canonical pair $\big(\lambda(\theta,\varphi), f(\theta,\varphi)\big)$ of boundary fields (equivalently: their spherical-harmonics components $(\lambda_{\ell m}, f_{\ell m})$) fields, which give rise to an additional term in the Hamiltonian (Gervais & Zwanziger 1980):
$$\sum_{\ell m} \phi_{\ell m}\big(E_{\ell m} - f_{\ell m}\big)$$
where $\phi_{\ell m}$ and $E_{\ell m}$ are the spherical-harmonics components of the scalar potential and the electric flux-density, $(\vec{n} \cdot \vec{E})$, on the boundary.
- In addition to Maxwell's equations in the bulk we get on the boundary:
$$\dot{\lambda}_{\ell m} = -\phi_{\ell m}\,, \qquad \dot{f}_{\ell m} = 0\,, \qquad E_{\ell m} = f_{\ell m}$$
where the third equation is the boundary part of Gauß' law.

Asymptotic DOF (contd.)

- **In the presence of charged states, the necessity to include degrees of freedom on the boundary does not disappear if the boundary is pushed to infinity.**
- A SSR for electric charge arises only if the conjugate 'position' observable, λ_{00}, is removed from the observables by **supplementary conditions**, like locality.
- In local QED there seems to arise an abundance of SSRs, not just that connected with the overall electric charge, but also for each asymptotic flux distribution.
- In the retarded representation of the EM fields, each asymptotic flux distribution is connected with incoming particle momenta. In order to avoid SSRs for incoming momenta—so as to be able to form incoming wave packets—we have to appropriately dress the incoming particle states with incoming radiation fields so as to comply with fixed asymptotic flux conditions. (Zwanziger 1976, Haller 1978, Fröhlich et. al 1979, Gervais & Zwanziger 1980, Buchholz 1982 ...)

Summary

- Derivations of so-called hard SSR usually rest on assumptions of exact symmetry and/or locality & causality. But whereas the derivations are (mathematically) exact, the hypotheses on which they rest are to be considered as physical idealisations of approximate validity.
- Derivations of so-called soft SSR are themselves subject to approximations right from the beginning. But the hypotheses on which they rest are more realistic.

> "The theoretical results currently available fall into two categories: rigorous results on approximate models and approximate results on realistic models"
>
> *A.S. Wightman & N. Glance 1989*